# Part III — Ramsey Theory

## Theorems with proof

Based on lectures by B. P. Narayanan

Notes taken by Dexter Chua

Lent 2017



What happens when we cut up a mathematical structure into many 'pieces' ? How big must the original structure be in order to guarantee that at least one of the pieces has a specific property of interest? These are the kinds of questions at the heart of Ramsey theory. A prototypical result in the area is van der Waerden's theorem, which states that whenever we partition the natural numbers into finitely many classes, there is a class that contains arbitrarily long arithmetic progressions.

The course will cover both classical material and more recent developments in the subject. Some of the classical results that I shall cover include Ramsey's theorem, van der Waerden's theorem and the Hales–Jewett theorem; I shall discuss some applications of these results as well. More recent developments that I hope to cover include the properties of non-Ramsey graphs, topics in geometric Ramsey theory, and finally, connections between Ramsey theory and topological dynamics. I will also indicate a number of open problems.

### Pre-requisites

There are almost no pre-requisites and the material presented in this course will, by and large, be self-contained. However, students familiar with the basic notions of graph theory and point-set topology are bound to find the course easier.

# Contents

**0 Introduction** **3**

**1 Graph Ramsey theory** **4**
1.1 Infinite graphs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
1.2 Finite graphs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7

**2 Ramsey theory on the integers** **12**

**3 Partition Regularity** **17**

**4 Topological Dynamics in Ramsey Theory** **23**

**5 Sums and products*** **30**

# 0 Introduction

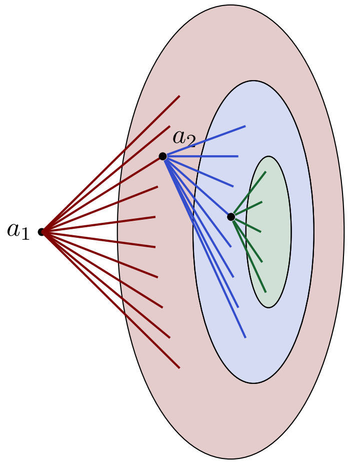
# 1 Graph Ramsey theory

## 1.1 Infinite graphs

**Theorem** (Ramsey's theorem)**.** Whenever we $k$-colour $\mathbb{N}^{(2)}$, there exists an infinite monochromatic set $X$, i.e. given any map $c : \mathbb{N}^{(2)} \to [k]$, there exists a subset $X \subseteq \mathbb{N}$ such that $X$ is infinite and $c|_{X^{(2)}}$ is a constant function.

*Proof.* Pick an arbitrary point $a_1 \in \mathbb{N}$. Then by the pigeonhole principle, there must exist an infinite set $\mathcal{B}_1 \subseteq \mathbb{N} \setminus \{a_1\}$ such that all the $a_1$-$\mathcal{B}_1$ edges (i.e. edges of the form $(a_1, b_1)$ with $b_1 \in \mathcal{B}_1$) are of the same colour $c_1$.

Now again arbitrarily pick an element $a_2 \in \mathcal{B}_1$. Again, we find some infinite $\mathcal{B}_2 \subseteq \mathcal{B}_1$ such that all $a_2$-$\mathcal{B}_2$ edges are the same colour $c_2$. We proceed inductively.




We obtain a sequence $\{a_1, a_2, \cdots\}$ and a sequence of colours $\{c_1, c_2, \cdots\}$ such that $c(a_i, a_j) = c_i$, for $i < j$.

Now again by the pigeonhole principle, since there are finitely many colours, there exists an infinite subsequence $c_{i_1}, c_{i_2}, \cdots$ that is constant. Then $a_{i_1}, a_{i_2}, \cdots$ is an infinite monochromatic set, since all edges are of the colour $c_{i_1} = c_{i_2} = \cdots$. So we are done. □

**Corollary** (Bolzano-Weierstrass theorem)**.** Let $(x_i)_{i \geq 0}$ be a bounded sequence of real numbers. Then it has a convergent subsequence.

*Proof.* We define a colouring $c : \mathbb{N}^{(2)} \to \{\uparrow, \downarrow\}$, where

$$c(ij) = \begin{cases} \uparrow & x_i < x_j \\ \downarrow & x_j \leq x_i \end{cases}$$

Then Ramsey's theorem gives us an infinite monochromatic set, which is the same as a monotone subsequence. Since this is bounded, it must convergence. □

**Theorem** (Ramsey's theorem for $r$ sets)**.** Whenever $\mathbb{N}^{(r)}$ is $k$-coloured, there exists an infinite monochromatic set, i.e. for any $c : \mathbb{N}^{(r)} \to [k]$, there exists an infinite $X \subseteq \mathbb{N}$ such that $c|_{X^{(r)}}$ is constant.

*Proof.* We induct on $r$. This is trivial when $r = 1$. Assume $r > 1$. We fix $a_1 \in \mathbb{N}$. We induce a $k$-colouring $c_1$ of $(\mathbb{N} \setminus \{a_1\})^{(r-1)}$ by

$$c_1(F) = c(F \cup \{a_1\}).$$

By induction, there exists an infinite $B_1 \subseteq \mathbb{N}\setminus\{a_1\}$ such that $B_1$ is monochromatic for $c_1$, i.e. all $a_1$-$B_1$ $r$-sets have the same colour $c_1$.

We proceed inductively as before. We get $a_1, a_2, a_3, \cdots$ and colours $c_1, c_2, \cdots$ etc. such that for any $r$-set $F$ contained in $\{a_1, a_2, \cdots\}$, we have $c(F) = c_{\min F}$.

Then again, there exists $c_{i_1}, c_{i_2}, c_{i_3}, \cdots$ all identical, and our monochromatic set is $\{a_{i_1}, a_{i_2}, a_{i_3}, \cdots\}$. □

**Theorem** (Canonical Ramsey theorem)**.** For any $c : \mathbb{N}^{(2)} \to \mathbb{N}$, there exists an infinite $X \subseteq \mathbb{N}$ such that one of the following hold:

(i) $c|_{X^{(2)}}$ is constant.

(ii) $c|_{X^{(2)}}$ is injective.

(iii) $c(ij) = c(k\ell)$ iff $i = k$ for all $i, j, k, \ell \in X$.

(iv) $c(ij) = c(k\ell)$ iff $j = \ell$ for all $i, j, k, \ell \in X$.

*Proof.* Consider the following colouring of $X^{(4)}$: let $c_1$ be a 2-colouring

$$c_1(ijk\ell) = \begin{cases} \texttt{SAME} & c(ij) = c(k\ell) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$

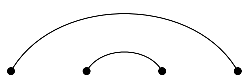
Then we know there is some infinite monochromatic set $X_1 \subseteq \mathbb{N}$ for $c_1$. If $X_1$ is coloured `SAME`, then we are done. Indeed, for any pair $ij$ and $i'j'$ in $X_1$, we can pick some huge $k, \ell$ such that $j, j' < k < \ell$, and then

$$c(ij) = c(k\ell) = c(i'j')$$

as we know $c_1(ijk\ell) = c_1(i'j'k\ell) = \texttt{SAME}$.

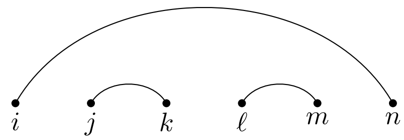
What if $X_1$ is coloured `DIFF`? We next look at what happens when we have edges that are nested each other. We define $c_2 : X_1^{(4)} \to \{\texttt{SAME}, \texttt{DIFF}\}$ defined by

$$c_2(ijk\ell) = \begin{cases} \texttt{SAME} & c(i\ell) = c(jk) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$



Again, we can find an infinite monochromatic subset $X_2 \subseteq X_1$ with respect to $c_2$.

We now note that $X_2$ cannot be coloured `SAME`. Indeed, we can just look at



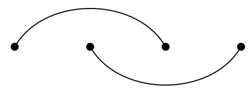

So if $X_2$ were SAME, we would have

$$c(\ell m) = c(in) = c(jk),$$

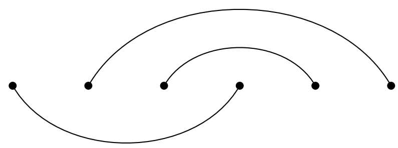
which is impossible since $X_1$ is coloured DIFF under $c_1$.

So $X_2$ is DIFF. Now consider $c_3 : X_2^{(4)} \to \{\texttt{SAME}, \texttt{DIFF}\}$ given by

$$c_3(ijk\ell) = \begin{cases} \texttt{SAME} & c(ik) = c(j\ell) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$



Again find an infinite monochromatic subset $X_3 \subseteq X_2$ for $c_3$. Then $X_3$ cannot be SAME, this time using the following picture:



contradicting the fact that $c_2$ is DIFF. So we know $X_3$ is DIFF.

We have now have ended up in this set $X_3$ such that if we have any two pairs of edges with different end points, then they must be different.

We now want to look at cases where things share a vertex. Consider $c_4 : X_3^{(3)} \to \{\texttt{SAME}, \texttt{DIFF}\}$ given by

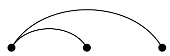
$$c_4(ijk) = \begin{cases} \texttt{SAME} & c(ij) = c(jk) \\ \texttt{DIFF} & \text{otherwise} \end{cases}$$

Let $X_4 \subseteq X_3$ be an infinite monochromatic set for $c_4$. Now $X_4$ cannot be coloured SAME, using the following picture:

which contradicts the fact that $c_1$ is DIFF. So we know $X_4$ is also coloured DIFF under $c_4$.

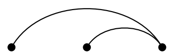
We are almost there. We need to deal with edges that nest in the sense of (iii) and (iv). We look at $c_5 : X_4^{(3)} \to \{\texttt{LSAME}, \texttt{LDIFF}\}$ given by

$$c_5(ijk) = \begin{cases} \texttt{LSAME} & c(ij) = c(ik) \\ \texttt{LDIFF} & \text{otherwise} \end{cases}$$



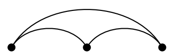
Again we find $X_5 \subseteq X_4$, an infinite monochromatic set for $c_5$. We don't separate into cases yet, because we know both cases are possible, but move on to classify the right case as well. Define $c_6 : X_5^{(3)} \to \{\texttt{RSAME}, \texttt{RDIFF}\}$ given by

$$c_5(ijk) = \begin{cases} \texttt{RSAME} & c(ik) = c(jk) \\ \texttt{RDIFF} & \text{otherwise} \end{cases}$$



Let $X_6 \subseteq X_5$ be an infinite monochromatic subset under $c_5$.

As before, we can check that it is impossible to get both LSAME and RSAME, using the following picture:



contradicting $c_4$ being DIFF.

Then the remaining cases (LDIFF, RDIFF), (LSAME, RDIFF) and (RDIFF, LSAME) corresponds to the cases (ii), (iii) and (iv). □

**Theorem** (Higher dimensional canonical Ramsey theorem)**.** Let $c : \mathbb{N}^{(r)} \to \mathbb{N}$ be a colouring. Then there exists $D \subseteq [r]$ and an infinite subset $X \subseteq \mathbb{N}$ such that for all $x, y \in X^{(r)}$, we have $c(x) = c(y)$ if $\{i : x_i = y_i\} \supseteq D$, where $x = \{x_1 < x_2 < \cdots < x_r\}$ (and similarly for $y$).

## 1.2 Finite graphs

**Theorem** (Finite Ramsey theorem)**.** For all $n$, we have $R(n) < \infty$.

*Proof.* Suppose not. Let $n$ be such that $R(n) = \infty$. Then for any $m \geq n$, there is a 2-colouring $c_m$ of $K_m = [m]^{(2)}$ such that there is no monochromatic set of size $n$.

We want to use these colourings to build up a colouring of $\mathbb{N}^{(2)}$ with no monochromatic set of size $n$. We want to say we take the "limit" of these colourings, but what does this mean? To do so, we need these colourings to be nested.

By the pigeonhole principle, there exists an infinite set $M_1 \subseteq \mathbb{N}$ and some fixed 2-colouring $d_n$ of $[n]$ such that $c_m|_{[n]^{(2)}} = d_n$ for all $m \in M_1$.

Similarly, there exists an infinite $M_2 \subseteq M_1$ such that $c_m|_{[n+1]^{(2)}} = d_{n+1}$ for $m \in M_2$, again for some 2-colouring $d_{n+1}$ of $[n+1]$. We repeat this over and over again. Then we get a sequence $d_n, d_{n+1}, \cdots$ of colourings such that $d_i$ is a 2-colouring of $[i]^{(2)}$ without a monochromatic $K_n$, and further for $i < j$, we have

$$d_j|_{[i]^{(2)}} = d_i.$$

We then define a 2-colouring $c$ of $\mathbb{N}^{(2)}$ by

$$c(ij) = d_m(ij)$$

for any $m > i, j$. Clearly, there exists no monochromatic $K_n$ in $c$, as any $K_n$ is finite. This massively contradicts the infinite Ramsey theorem. □

*Proof.* Suppose $R(n) = \infty$. Consider the theory with proposition letters $p_{ij}$ for each $ij \in \mathbb{N}^{(2)}$. We will think of the edge $ij$ as red if $p_{ij} = \top$, and blue if $p_{ij} = \bot$. For each subset of $\mathbb{N}$ of size $n$, we add in the axiom that says that set is not monochromatic.

Then given any finite subset of the axioms, it mentions only finitely many subsets of $\mathbb{N}$. Suppose it mentions vertices only up to $m \in \mathbb{N}$. Then by assumption, there is a 2-colouring of $[m]^{(2)}$ with no monochromatic subset of size $n$. So by assigning $p_{ij}$ accordingly (and randomly assigning the remaining ones), we have found a model of this finite subtheory. Thus every finite fragment of the theory is consistent. Hence the original theory is consistent. So there is a model, i.e. a colouring of $\mathbb{N}^{(2)}$ with no monochromatic subset.

This contradicts the infinite Ramsey theorem. □

**Theorem.** We have

$$R(n, m) \leq R(n-1, m) + R(n, m-1).$$

for all $n, m \in \mathbb{N}$. Consequently, we have

$$R(n, m) \leq \binom{n+m-1}{n-2}.$$

*Proof.* We induct on $n + m$. It is clear that

$$R(1, n) = R(n, 1) = 1, \quad R(n, 2) = R(2, n) = n$$

Now suppose $N \geq R(n-1, m) + R(n, m-1)$. Consider any red-blue colouring of $K_N$ and any vertex $v \in V(K_n)$. We write

$$v(K_n) \setminus \{v\} = A \cup B,$$

where each vertex $A$ is joined by a red edge to $v$, and each vertex in $B$ is joined by a blue edge to $v$. Then

$$|A| + |B| \geq N - 1 \geq R(n-1, m) + R(n, m-1) - 1.$$

It follows that either $|A| \geq R(n-1, m)$ or $|B| \geq R(n, m-1)$. We wlog it is the former. Then by definition of $R(n-1, m)$, we know $A$ contains either a blue copy of $K_m$ or a red copy of $K_{n-1}$. In the first case, we are done, and in the second case, we just add $v$ to the red $K_{n-1}$.

The last formula is just a straightforward property of binomial coefficients. □

**Theorem.** We have $R(n) \geq \sqrt{2}^n$ for sufficiently large $n \in \mathbb{N}$.

*Proof.* Let $N \leq \sqrt{2}^n$. We perform a red-blue colouring of $K_N$ randomly, where each edge is coloured red independently of the others with probability $\frac{1}{2}$.

We let $X_R$ be the number of red copies of $K_n$ in such a colouring. Then since expectation is linear, we know the expected value is

$$\begin{aligned}[X_R] &= \binom{N}{n}\left(\frac{1}{2}\right)^{\binom{n}{2}} \\ &\leq \left(\frac{eN}{n}\right)^n \left(\frac{1}{2}\right)^{\binom{n}{2}} \\ &\leq \left(\frac{e}{n}\right)^n \sqrt{2}^{n^2} \sqrt{2}^{-n^2} \\ &= \left(\frac{e}{n}\right)^n \\ &< \frac{1}{2}\end{aligned}$$

for sufficiently large $n$.

Similarly, we have $[X_B] < \frac{1}{2}$. So the expected number of monochromatic $K_n$ is $< 1$. So in particular there must be some colouring with no monochromatic $K_n$. □

**Theorem.** We have

$$R(3, n) \leq \frac{100n^2}{\log n}$$

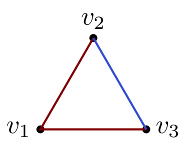
for sufficiently large $n \in \mathbb{N}$.

*Proof.* Let $N \geq 100n^2/\log n$, and consider a red-blue colouring of the edges of $K_N$ with no red $K_3$. We want to find a blue $K_n$ in such a colouring.

We may assume that

(i) No vertex $v$ has $\geq n$ red edges incident to it, as argued just now.

(ii) If we have $v_1, v_2, v_3$ such that $v_1v_2$ and $v_1v_3$ are red, then $v_2v_3$ is blue:




We now let

$$\mathcal{F} = \{W : W \subseteq V(K_N) \text{ such that } W^{(2)} \text{ is monochromatic and blue}\}.$$

We want to find some $W \in \mathcal{F}$ such that $|W| \geq n$, i.e. a blue $K_n$. How can we go about finding this?

Let $W$ be a uniformly random member of $\mathcal{F}$. We will be done if we can show that that $\mathbb{E}[|W|] \geq n$.

We are going to define a bunch of random variables. For each vertex $v \in V(K_N)$, we define the variable

$$X_v = n\mathbf{1}_{\{v \in W\}} + |\{u : uv \text{ is red and } u \in W\}|.$$

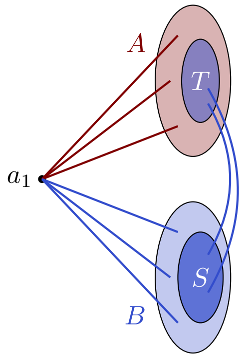
**Claim.**

$$\mathbb{E}[X_v] > \frac{\log n}{10}$$

for each vertex $v$.

To see this, let
$$A = \{u : uv \text{ is red}\}$$
and let
$$B = \{u : uv \text{ is blue}\}.$$
then from the properties we've noted down, we know that $|A| < n$ and $A^{(2)}$ is blue. So we know very well what is happening in $A$, and nothing about what is in $B$.

We fix a set $S \subseteq B$ such that $S \in \mathcal{F}$, i.e. $S^{(2)}$ is blue. What can we say about $W$ if we condition on $B \cap W = S$?




Let $T \subseteq A$ be the set of vertices that are joined to $S$ only by blue edges. Write $|T| = x$. Then if $B \cap W = S$, then either $W \subseteq S \cup T$, or $W \subseteq S \cup \{v\}$. So there are exactly $2^x + 1$ choices of $W$. So we know

$$\begin{aligned}\mathbb{E}[X_v \mid W \cap B = S] &\geq \frac{n}{2^x+1} + \frac{2^x}{2^x+1}(\mathbb{E}[|\text{random subset of T}|])\\ &= \frac{n}{2^x+1} + \frac{2^{x-1}x}{2^x+1}.\end{aligned}$$

Now if
$$x < \frac{\log n}{2},$$
then
$$\frac{n}{2^x+1} \geq \frac{n}{\sqrt{n}+1} \geq \frac{\log n}{10}$$
for all sufficiently large $n$. On the other hand, if
$$x \geq \frac{\log n}{2},$$
then
$$\frac{2^{x-1}x}{2^x+1} \geq \frac{1}{4} \cdot \frac{\log n}{2} \geq \frac{\log n}{10}.$$
So we are done.

**Claim.**
$$\sum_{v \in V} X_v \leq 2n|W|.$$

To see this, for each vertex $v$, we know that if $v \in W$, then it contributes $n$ to the sum via the first term. Also, by our initial observation, we know that $v \in W$ has at most $n$ neighbours. So it contributes at most $n$ to the second term (acting as the "$u$").

Finally, we know that

$$\mathbb{E}[|W|] \geq \frac{1}{2n} \sum \mathbb{E}[X_V] \geq \frac{N}{2n} \frac{\log n}{10} \geq \frac{100n^2}{\log n} \cdot \frac{\log n}{20n} \geq 5n.$$

Therefore we can always find some $W \in \mathcal{F}$ such that $|W| \geq n$. $\square$

# 2 Ramsey theory on the integers

**Theorem** (van der Waerden theorem)**.** Let $m, k \in N$. Then there is some $N = W(m, k)$ such that whenever $[N]$ is $k$-coloured, then there is a monochromatic arithmetic progression of length $n$.

*Proof.* We induct on $m$. The result is clearly trivial when $m = 1$, and follows easily from the pigeonhole principle when $m = 2$.

Suppose $m > 1$, and assume inductively that $W(m-1, k')$ exists for any $k' \in \mathbb{N}$.

Here is the claim we are trying to established:

**Claim.** For each $r \leq k$, there is a natural number $n$ such that whenever we $k$-colour $[n]$, then either

(i) There exists a monochromatic AP of length $m$; or

(ii) There are $r$ colour-focused AP's of length $m - 1$.

It is clear that this claim implies the theorem, as we can pick $r = k$. Then if there isn't a monochromatic AP of length $m$, then we look at the colour of the common focus, and it must be one of the colours of those AP's.

To prove the claim, we induct on $r$. When $n = 1$, we may take $W(m-1, k)$. Now suppose $r > 1$, and some $n'$ works for $r - 1$. With the benefit of hindsight, we shall show that

$$n = W(m - 1, k^{2n'})2n'$$

works for $r$.

We consider any $k$-colouring of $[n]$, and suppose it has no monochromatic AP of length $m$. We need to find $r$ colour-focused progressions of length $n - 1$.

We view this $k$-colouring of $[n]$ as a $k^{2n'}$ colouring of blocks of length $2n'$, of which there are $W(m - 1, k^{2n'})$.

Then by definition of $W(m - 1, k^{2n'})$, we can find blocks

$$B_s, B_{s+t}, \cdots, B_{s+(m-2)t}$$

which are coloured identically. By the inductive hypothesis, we know each $B_s$ contains $r - 1$ colour-focused AP's of length $m - 1$, say $A_1, .., A_{r-1}$ with first terms $a_1, \cdots, a_r$ and common difference $d_1, \cdots, d_{r-1}$, and also their focus $f$, because the length of $B_s$ is $2n'$, not just $n'$.

Since we assumed there is no monochromatic progression of length $n$, we can assume $f$ has a different colour than the $A_i$.

Now consider $A'_1, A'_2, \cdots, A'_{r-1}$, where $A'_i$ has first term $a_i$, common difference $d_i + 2n't$, and length $m - 1$. This difference sends us to the next block, and then the next term in the AP. We also pick $A'_r$ to consist of the all the focus of the blocks $B_i$, namely

$$A'_r = \{f, f + 2n't, \cdots, f + 2n't(m - 2)\}$$

These progressions are monochromatic with distinct colours, and focused at $f + (2n't)(m - 1)$. So we are done. □

**Theorem** (Gallai)**.** Whenever $\mathbb{N}^d$ is $k$-coloured, there exists a monochromatic (homothetic) copy of $S$ for each finite $S \subseteq \mathbb{N}^d$.

**Theorem** (Hales–Jewett theorem)**.** For all $m, k \in \mathbb{N}$, there exists $N = HJ(m, k)$ such that whenever $[m]^N$ is $k$-coloured, there exists a monochromatic line.

*Proof.* We proceed by induction on $m$. This is clearly trivial for $m = 1$, as a line only has a single point.

Now suppose $m > 1$, and that $HJ(m-1, k)$ exists for all $k \in \mathbb{N}$. As before, we will prove the following claim:

**Claim.** For each $1 \leq r \leq k$, there is an $n \in \mathbb{N}$ such that in any $k$-colouring of $[m]^n$, either

(i) there exists a monochromatic line; or

(ii) there exists $r$ colour-focused lines.

Again, the result is immediate from the claim, as we just use it for $r = k$ and look at the colour of the focus.

The prove this claim, we induct on $r$. If $r = 1$, then picking $n = HJ(m-1, k)$ works, as a single colour-focused line of length $n$ is just a monochromatic line of length $n - 1$, and $[m-1]^n \subseteq [m]^n$ naturally.

Now suppose $r > 1$ and $n$ works for $r - 1$. Then we will show that $n + n'$ works, where

$$n' = HJ(m-1, k^{m^n}).$$

Consider a colouring $c : [m]^{n+n'} \to [k]$, and we assume this has no monochromatic lines.

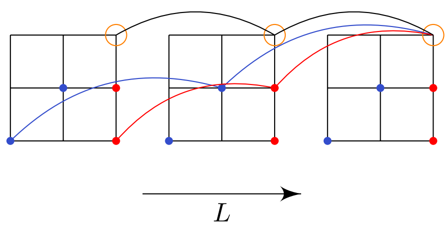
We think of $[m]^{n+n'}$ as $[m]^n \times [m]^{n'}$. So for each point in $[m]^{n'}$, we have a whole cube $[m]^n$. Consider a $k^{m^n}$ colouring of $[m]^{n'}$ as follows — given any $x \in [m]^{n'}$, we look at the subset of $[m]^{n+n'}$ containing the points with last $n'$ coordinates $x$. Then we define the new colouring of $[m]^{n'}$ to be the restriction of $c$ to this $[m]^n$, and there are $m^n$ possibilities.

Now there exists a line $L$ such that $L \setminus L^+$ is monochromatic for the new colouring. This means for all $a \in [m]^n$ and for all $b, b' \in L \setminus L^+$, we have

$$c(a, b) = c(a, b').$$

Let $c''(a)$ denote this common colour for all $a \in [m]^n$. this is a $k$-colouring of $[m]^n$ with no monochromatic line (because $c$ doesn't have any). So by definition of $n$, there exists lines $L_1, L_2, \cdots, L_{r-1}$ in $[m]^n$ which are colour-focused at some $f \in [m]^n$ for $c''$.

In the proof of van der Waerden, we had a progression within each block, and also how we just between blocks. Here we have the same thing. We have the lines in $[m]^n$, and also the "external" line $L$.



Consider the line $L'_1, L'_2, \cdots, L'_{r-1}$ in $[m]^{n+n'}$, where

$$(L'_i)^- = (L_i^-, L^-),$$

and the active coordinate set is $I_i \cup I$, where $I$ is the active coordinate set of $L$.

Also consider the line $F$ with $F^- = (f, L^-)$ and active coordinate set $I$.

Then we see that $L'_1, \cdots, L'_{r-1}, F$ are $r$ colour-focused lines with focus $(f, L^+)$. □

**Theorem** (Gallai)**.** Whenever $\mathbb{N}^d$ is $k$-coloured, there exists a monochromatic (homothetic) copy of $S$ for each finite $S \subseteq \mathbb{N}^d$.

*Proof.* Let $S = \{S(1), S(2), \cdots, S(m)\} \subseteq \mathbb{N}^d$. Given a $k$-colouring $\mathbb{N}^d \to [k]$, we induce a $k$-colouring $c : [m]^N \to [k]$ by

$$c'(x_1, \cdots, x_N) = c(S(x_1) + S(x_2) + \cdots + S(x_N)).$$

By Hales-Jewett, for sufficiently large $N$, there exists a monochromatic line $L$ for $c'$, which gives us a monochromatic homothetic copy of $S$ in $\mathbb{N}^d$. For example, if the line is (1 2 1), (2 2 2) and (3 2 3), then we know

$$c(S(1) + S(2) + S(1)) = c(S(2) + S(2) + S(2)) = c(S(3) + S(2) + S(3)).$$

So we have the monochromatic homothetic copy $\lambda S + \mu$ , where $\lambda = 2$ (the number of active coordinates), and $\mu = S(2)$. □

**Theorem** (Szemerédi theorem)**.** Let $\delta > 0$ and $m \in \mathbb{N}$. Then there exists some $N = S(m, \delta) \in \mathbb{N}$ such that any subset $A \subseteq [N]$ with $|A| \geq \delta N$ contains an $m$-term arithmetic progression.

**Theorem.** For any $c : \mathbb{N} \to \mathbb{N}$, there exists a $m$-AP on which either

(i) $c$ is constant; or

(ii) $c$ is injective.

*Proof.* We set

$$\delta = \frac{1}{m^2(m+1)^2}.$$

We let $N = S(m, \delta)$. We write

$$[N] = A_1 \cup A_2 \cup \cdots \cup A_k,$$

where the $A_i$'s are the colour-classes of $c|_{[N]}$. By choice of $N$, we are done if $|A_i| \geq \delta N$ for some $1 \leq i \leq k$. So suppose not.

Let's try to count the number of arithmetic progressions in $[N]$. There are more than $N^2/(m+1)^2$ of these, as we can take any $a, d \in [N/m + 1]$. We want to show that there is an AP that hits each $A_i$ at most once.

So, fixing an $i$, how many AP's are there that hit $A_i$ at least twice? We need to pick two terms in $A_i$, and also decide which two terms in the progression they are in, e.g. they can be the first and second term, or the 5th and 17th term. So there are at most $m^2|A_i|^2$ terms.

So the number of AP's on which $c$ is injective is greater than

$$\frac{N^2}{(m+1)^2} - k|A_i|^2 m^2 \geq \frac{N^2}{(m+1)^2} - \frac{1}{\delta}(\delta N)^2 m^2 = \frac{N^2}{(m+1)^2} - \delta N^2 m^2 \geq 0.$$

So we are done. Here the first inequality follows from the fact that $\sum |A_i| = [N]$ and each $|A_i| < \delta N$. □

**Theorem.** For any $c : \mathbb{N}^{(2)} \to \{\text{red}, \text{blue}\}$, either

(i) There exists a blue $m$-AP for each $m \in \mathbb{N}$; or

(ii) There exists arbitrarily large red sets.

*Proof.* Suppose we can't find a blue $m$-AP for some fixed $m$. We induct on $r$, and try to find a red set of size $r$.

Say $A \subseteq \mathbb{N}$ is a progression of length $N$. Since $A$ has no blue $m$-term progression, so it must contain many red edges. Indeed, each $m$-AP in $A$ must contain a red edge. Also each edge specifies two points, and this can be extended to an $m$-term progression in at most $m^2$ ways. Since there are $N^2/(m+1)^2$. So there are at least

$$\frac{N^2}{m^2(m+1)^2}$$

red edges. With the benefit of hindsight, we set

$$\delta = \frac{1}{2m^2(m+1)^2}.$$

The idea is that since we have lots of red edges, we can try to find a point with a lot of red edges connected to it, and we hope to find a progression in it.

We say $X, Y \subseteq \mathbb{N}$ form an $(r, k)$-structure if

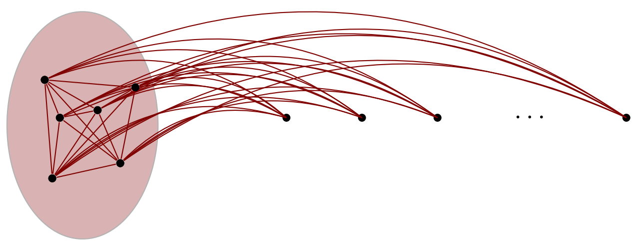
(i) They are disjoint

(ii) $X$ is red;

(iii) $Y$ is an arithmetic progression;

(iv) All $X$-$Y$ edges are red;

(v) $|X| = r$ and $|Y| = k$.



We show by induction that there is an $(r, k)$-structure for each $r$ and $k$.

A $(1, k)$ structure is just a vertex connected by red edges to a $k$-point structure. If we take $N = S(\delta, k)$, we know among the first $N$ natural numbers, there are at least $N^2/(m^2(m+1)^2)$ red edges inside $[N]$. So in particular, some $v \in [N]$ has at least $\delta N$ red neighbours in $[N]$, and so we know $v$ is connected to some $k$-AP by red edges. That's the base case done.

Now suppose we can find an $(r-1, k')$-structure for all $k' \in \mathbb{N}$. We set

$$k' = S\left(\frac{1}{2m^2(m+1)^2}, k\right).$$

We let $(X, Y)$ be an $(r-1, k')$-structure. As before, we can find $v \in Y$ such that $v$ has $\delta|Y|$ red neighbours in $Y$. Then we can find a progression $Y'$ of length $k$ in the red neighbourhood of $v$, and we are done, as $(X \cup \{v\}, Y')$ is an $(r, k)$-structure, and an "arithmetic progression" within an arithmetic progression is still an arithmetic progression. □

# 3 Partition Regularity

**Theorem** (Rado's theorem)**.** A matrix $A$ is partition regular iff it has the column property.

**Theorem.** If $a_1, \cdots, a_n \in \mathbb{Q} \setminus \{0\}$, then $\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is partition regular iff there exists a non-empty $I \subseteq [n]$ such that

$$\sum_{i \in I} a_i = 0.$$

**Proposition.** If $a_2, \cdots, a_n \in \mathbb{Q} \setminus \{0\}$ and $\begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}$ is partition regular, then

$$\sum_{i \in I} a_i = 0$$

for some non-empty $I$.

*Proof.* We wlog $a_i \in \mathbb{Z}$, by scaling. Fix a suitably large prime

$$p > \sum_{i=1}^n |a_i|,$$

and consider the $(p-1)$-colouring of $\mathbb{N}$ where $x$ is coloured $d(x)$. We find $x_1, \cdots, x_n$ such that

$$\sum a_i x_i = 0.$$

and $d(x_i) = d$ for some $d \in \{1, \cdots, p-1\}$. We write out everything in base $p$, and let

$$L = \min\{L(x_i) : 1 \leq i \leq n\},$$

and set

$$I = \{i : L(x_i) = L\}.$$

Then for all $i \in I$, we have

$$x_i \equiv d \pmod{p^{L+1}}.$$

On the other hand, we are given that

$$\sum a_i x_i = 0.$$

Taking mod $p^{L+1}$ gives us

$$\sum_{i \in I} a_i d = 0 \pmod{p^{L+1}}.$$

Since $d$ is invertible, we know

$$\sum_{i \in I} a_i = 0 \pmod{p^{L+1}}.$$

But by our choice of $p$, this implies that $\sum_{i \in I} a_i = 0$. ☐

**Proposition.** The equation $\begin{pmatrix} 1 & \lambda & -1 \end{pmatrix}$ is partition regular for all $\lambda \in \mathbb{Q}$.

*Proof.* We may wlog $\lambda > 0$. If $\lambda = 0$, then this is trivial, and if $\lambda < 0$, then we can multiply the whole equation by $-1$.

Say

$$\lambda = \frac{r}{s}.$$

The idea is to try to find solutions of this in long arithmetic progressions.

Suppose $\mathbb{N}$ is $k$-coloured. We let

$$\{a, a+d, \cdots, a+nd\}$$

be a monochromatic AP, for $n$ sufficiently large.

If $sd$ were the same colour as this AP, then we'd be done, as we can take

$$x = a, \quad y = sd, \quad z = a + \frac{r}{s}sd.$$

In fact, if any of $sd, 2sd, \cdots, \lfloor \frac{n}{r} \rfloor sd$ have the same colour as the AP, then we'd be done by taking

$$x = a, \quad y = isd, \quad z = a + \frac{r}{s}isd = a + ird \leq a + nd.$$

If this were not the case, then $\{sd, 2sd, \cdots, \lfloor \frac{n}{r} \rfloor sd\}$ is $(k-1)$-coloured, and this is just a scaled up copy of $\mathbb{N}$. So we are done by induction on $k$. □

**Theorem.** If $a_1, \cdots, a_n \in \mathbb{Q}$, then $\begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is partition regular iff there exists a non-empty $I \subseteq [n]$ such that

$$\sum_{i \in I} a_i = 0.$$

*Proof.* One direction was done. To do the other direction, we recall that we had a really easy case of, say,

$$\begin{pmatrix} 2 & 3 & -5 \end{pmatrix},$$

because we can just make all the variables the same?

In the general case, we can't quite do this, but we may try to solve this equation with the least number of variables possible. In fact, we shall find some monochromatic $x, y, z$, and then assign each of $x_1, \cdots, x_n$ to be one of $x, y, z$.

We know

$$\sum_{i \in I} a_i = 0.$$

We now partition $I$ into two pieces. We fix $i_0 \in I$, and set

$$x_i = \begin{cases} x & i = i_0 \\ y & i \in I \setminus \{i_0\} \\ z & i \notin I \end{cases}.$$

We would be done if whenever we finitely colour $\mathbb{N}$, we can find monochromatic $x, y, z$ such that

$$a_{i_0} x + \left(\sum_{i \in I \setminus \{i_0\}} a_i\right) z + \left(\sum_{i \notin I} a_i\right) y = 0.$$

But, since

$$\sum_{i \in I} a_i = 0,$$

this is equivalent to

$$a_{i_0} x - a_{i_0} z + (\text{something}) y = 0.$$

Since all these coefficients were non-zero, we can divide out by $a_{i_0}$, and we are done by the previous case. □

**Proposition.** If $A$ is an $m \times n$ matrix with rational entries which is partition regular, then $A$ has the columns property.

*Proof.* We again wlog all entries of $A$ are integers. Let the columns of $A$ be $c^{(1)}, \cdots, c^{(n)}$. Given a prime $p$, we consider the $(p-1)$-colouring of $\mathbb{N}$, where $x$ is coloured $d(x)$, the last non-zero digit in the base $p$ expansion. Since $A$ is partition regular, we obtain a monochromatic solution.

We then get a monochromatic $x_1, \cdots, x_n$ such that $Ax = 0$, i.e.

$$\sum x_i c^{(i)} = 0.$$

Any such solution with colour $d$ induces a partition of $[n] = B_1 \cup B_2 \cup \cdots \cup B_r$, where

- For all $i, j \in B_s$, we have $L(x_i) = L(x_j)$; and
- For all $s < t$ and $i \in B_s$, $j \in B_t$, the $L(x_i) < L(x_j)$.

Last time, with the benefit of hindsight, we were able to choose some large prime $p$ that made the argument work. So we use the trick we mentioned after the proof last time.

Since there are finitely many possible partitions of $[n]$, we may assume that this particular partition is generated by infinitely many primes $p$. Call these primes $\mathcal{P}$. We introduce some notation. We say two vectors $u, v \in \mathbb{Z}^m$ satisfy $u \equiv v \pmod{p}$ if $u_i \equiv v_i \pmod{p}$ for all $i = 1, \cdots, m$.

Now we know that

$$\sum x_i c^{(i)} = 0.$$

Looking at the first non-zero digit in the base $p$ expansion, we have

$$\sum_{i \in B_1} d c^{(i)} \equiv 0 \pmod{p}.$$

From this, we conclude that, by multiplying by $d^{-1}$

$$\sum_{i \in B_1} c^{(i)} \equiv 0 \pmod{p},$$

for all $p \in \mathcal{P}$. So we deduce that

$$\sum_{i \in B_1} c^{(i)} = 0.$$

Similarly, for higher $s$, we find that for each base $p$ colouring, we have

$$\sum_{i \in B_s} p^t d c^{(i)} + \sum_{i \in B_1 \cup \ldots \cup B_s} x_i c^{(i)} \equiv 0 \pmod{p^{t+1}}$$

for all $s \geq 2$, and some $t$ dependent on $s$ and $p$. Multiplying by $d^{-1}$, we find

$$\sum_{i \in B_s} p^t c^{(i)} + \sum_{i \in B_1 \cup \ldots \cup B_{s-1}} (d^{-1} x_i) c^{(i)} \equiv 0 \pmod{p^{t+1}}. \qquad (*)$$

We claim that this implies

$$\sum_{i \in \mathcal{B}_s} c^{(i)} \in \langle c^{(i)} : i \in B_1 \cup \cdots \cup B_{s-1} \rangle.$$

This is not exactly immediate, because the values of $x_i$ in $(*)$ may change as we change our $p$. But it is still some easy linear algebra.

Suppose this were not true. Since we are living in a Euclidean space, we have an inner product, and we can find some $v \in \mathbb{Z}^m$ such that

$$\langle v, c^{(i)} \rangle = 0 \text{ for all } i \in B_1 \cup \cdots \cup B_{s-1},$$

and

$$\left\langle v, \sum_{i \in B_s} c^{(i)} \right\rangle \neq 0.$$

But then, taking inner products with gives

$$p^t \left\langle v, \sum_{i \in B_s} c^{(i)} \right\rangle \equiv 0 \pmod{p^{t+1}}.$$

Equivalently, we have

$$\left\langle v, \sum_{i \in B_s} c^{(i)} \right\rangle \equiv 0 \pmod{p},$$

but this is a contradiction. So we showed that $A$ has the columns property with respect to the partition $[n] = B_1 \cup \cdots \cup B_r$. □

**Proposition.** Let $m, p, c \in \mathbb{N}$. Then whenever $\mathbb{N}$ is finitely coloured, there exists a monochromatic $(m, p, c)$-set.

*Proof.* It suffices to find an $(m, p, c)$-set all of whose rows are monochromatic, since when $\mathbb{N}$ is $k$-coloured, and $(m', p, c)$-set with $m' = mk + 1$ has $m$ monochromatic rows of the same colour by pigeonhole, and these rows contain a monochromatic $(m, p, c)$-set, by restricting to the elements where a lot of the $\lambda_i$ are zero. In this proof, whenever we say $(m, p, c)$-set, we mean one all of whose rows are monochromatic.

We will prove this by induction. We have a $k$-colouring of $[n]$, where $n$ is very very very large. This contains a $k$-colouring of

$$B = \left\{ c, 2c, \cdots, \left\lfloor \frac{n}{c} \right\rfloor c \right\}.$$

Since $c$ is fixed, we can pick this so that $\frac{n}{c}$ is large. By van der Waerden, we find some set monochromatic

$$A = \{cx_1 - Nd, cx_1 - (N-1)d, \cdots, cx_1 + Nd\} \subseteq B,$$

with $N$ very very large. Since each element is a multiple of $c$ by assumption, we know that $c \mid d$. By induction, we may find an $(m-1, p, c)$-set in the set $\{d, 2d, \cdots, Md\}$, where $M$ is large. We are now done by the $(m, p, c)$ set on generators $x_1, \cdots, x_n$, provided

$$cx_1 + \sum_{i=2}^{m} \lambda_i x_i \in A$$

for all $\lambda_i \in [-p, p]$, which is easily seen to be the case, provided $N \geq (m-1)pM$. □

**Corollary** (Finite sum theorem)**.** For every fixed $m$, whenever we finitely-colour $\mathbb{N}$, there exists $x_1, \cdots, x_m$ such that $FS(x_1, \cdots, x_m)$ is monochromatic.

**Proposition.** If $A$ is a rational matrix with the columns property, then there is some $m, p, c \in \mathbb{N}$ such that $Ax = 0$ has a solution inside any $(m, p, c)$ set, i.e. all entries of the solution lie in the $(m, p, c)$ set.

*Proof.* We again write

$$A = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ c^{(1)} & c^{(2)} & \cdots & c^{(n)} \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}.$$

Re-ordering the columns of $A$ if necessary, we assume that we have

$$[n] = B_1 \cup \cdots \cup B_r$$

such that $\max(B_s) < \min(B_{s+1})$ for all $s$, and we have

$$\sum_{i \in B_s} c^{(i)} = \sum_{i \in B_1 \cup \ldots \cup B_{s-1}} q_{is} c^{(i)}$$

for some $q_{is} \in \mathbb{Q}$. These $q_{is}$ only depend on the matrix. In other words, we have

$$\sum d_{is} c^{(i)} = 0,$$

where

$$d_{is} = \begin{cases} -q_{is} & i \in B_1 \cup \cdots B_{s-1} \\ 1 & i \in B_s \\ 0 & \text{otherwise} \end{cases}$$

For a fixed $s$, if we scan these coefficients starting from $i = n$ and then keep decreasing $i$, then the first non-zero coefficient we see is 1, which is good, because it looks like what we see in an $(m, p, c)$ set.

Now we try to write down a general solution with $r$ many free variables. Given $x_1, \cdots, x_r \in \mathbb{N}^r$, we look at

$$y_i = \sum_{s=1}^{r} d_{is} x_s.$$

It is easy to check that $Ay = 0$ since

$$\sum y_i c^{(i)} = \sum_i \sum_s d_{is} x_s c^{(i)} = \sum_s x_s \sum_i d_{is} c^{(i)} = 0.$$

Now take $m = r$, and pick $c$ large enough such that $cd_{is} \in \mathbb{Z}$ for all $i, s$, and finally, $p = \max\{cd_{is} : i, s \in \mathbb{Q}\}$.

Thus, if we consider the $(m, p, c)$-set on generators $(x_m, \cdots, x_1)$ and $y_i$ as defined above, then we have $Ay = 0$ and hence $A(cy) = 0$. Since $cy$ is integral, and lies in the $(m, p, c)$ set, we are done! □

**Theorem** (Rado's theorem)**.** A matrix $A$ is partition regular iff it has the column property.

**Corollary** (Consistency theorem)**.** If $A$, $B$ are partition regular in independent variables, then

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

is partition regular. In other words, we can solve $Ax = 0$ and $By = 0$ simultaneously in the same colour class.

*Proof.* The matrix

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

has the columns property if $A$ and $B$ do. □

**Corollary.** Whenever $\mathbb{N}$ is finitely-coloured, one colour class contains solutions to *all* partition regular systems!

*Proof.* Suppose not. Then we have $\mathbb{N} = D_1 \cup \cdots \cup D_k$ such that for each $D_i$, there is some partition regular matrix $A_i$ such that we cannot solve $A_i x = 0$ inside $D_i$. But this contradicts the fact that $\operatorname{diag}(A_1, A_2, \cdots, A_k)$ is partition regular (by applying consistency theorem many times). □

# 4 Topological Dynamics in Ramsey Theory

**Theorem** (Topological van der Waerden)**.** Let $(X, T)$ be a dynamical system. Then there exists an $\varepsilon > 0$ such that whenever $r \in \mathbb{N}$, then we can find $x \in X$ and $n \in \mathbb{N}$ such that $\rho(x, T^{in}x) < \varepsilon$ for all $i = 1, \cdots, r$.

**Proposition.** $(\bar{X}, T)$ is a dynamical system.

*Proof.* It suffices to show that $\bar{x}$ is closed under $T$. If $y \in \bar{x}$, then we have some $s_i$ such that $T^{s_i}x \to y$. Since $T$ is continuous, we know $T^{s_i+1}x \to Ty$. So $Ty \in \bar{x}$. Similarly, $T^{-1}\bar{x} = \bar{x}$. □

**Corollary** (van der Waerden theorem)**.** Let $r, k \in \mathbb{N}$. Then whenever $\mathbb{Z}$ is $k$-coloured, then there is a monochromatic arithmetic progression of length $r$.

*Proof.* Let $c : \mathbb{Z} \to [k]$. Consider $(\bar{c}, \mathcal{L})$. By topological van der Waerden, we can find $x \in \bar{c}$ and $n \in \mathbb{N}$ such that $\rho(x, \mathcal{L}^{in}x) < 1$ for all $i = 1, \cdots, r$. In particular, we know that $x$ and $\mathcal{L}^{in}x$ agree at 0. So we know that

$$x(0) = \mathcal{L}^{in}x(0) = x(in)$$

for all $i = 0, \cdots, r$.

We are not quite done, because we just know that $x \in \bar{c}$, and not that $x = \mathcal{L}^k x$ for some $k$.

But this is not bad. Since $x \in \bar{c}$, we can find some $s \in \mathbb{Z}$ such that

$$\rho(T^s c, x) \leq \frac{1}{rn + 1}.$$

This means $x$ and $T^s c$ agree on the first $rn + 1$ elements. So we know

$$c(s) = c(s + n) = \cdots = c(s + rn).$$

□

**Proposition.** We have $c' \in \bar{c}$ iff $\operatorname{Seq}(c') \subseteq \operatorname{Seq}(c)$.

*Proof.* We first prove $\Rightarrow$. Suppose $c' \in \bar{c}$. Let $(c'(i), \cdots, c'(i + r - 1)) \in \operatorname{Seq}(c)$. Then we have $s \in \mathbb{Z}$ such that

$$\rho(c', \mathcal{L}^s c) < \frac{1}{1 + \max(|i|, |i + s - 1|)},$$

which implies

$$(c(s + i), \cdots, c(s + i + r - 1)) = (c'(i), \cdots, c'(i + s - 1)).$$

So we are done.

For $\Leftarrow$, if $\operatorname{Seq}(c') \subseteq \operatorname{Seq}(c)$, then for all $n \in \mathbb{N}$, there exists $s_n \in \mathbb{Z}$ such that

$$(c'(-n), \cdots, c'(n)) = (c(s_n - n), \cdots, c(s_n + n)).$$

Then we have

$$\mathcal{L}^{s_i} c \to c'.$$

So we have $c' \in \bar{c}$. □

**Proposition.** Every dynamical system $(X, T)$ has a minimal point.

*Proof.* Let $\mathcal{U} = \{\bar{x} : x \in X\}$. Thus is a family of closed sets, ordered by inclusion. We want to apply Zorn's lemma to obtain a minimal element. Consider a chain $S$ in $\mathcal{U}$. If

$$\bar{x}_1 \supseteq \bar{x}_2 \supseteq \cdots \supseteq \bar{x}_n,$$

then their intersection is $\bar{x}_n$, which is in particular non-empty. So any finite collection in $S$ has non-empty intersection. Since $X$ is compact, we know

$$\bigcap_{\bar{x} \in S} \bar{x} \neq \emptyset.$$

We pick

$$z \in \bigcap_{\bar{x} \in S} \bar{x} \neq \emptyset.$$

Then we know that

$$\bar{z} \subseteq \bar{x}$$

for all $\bar{x} \in S$. So we know

$$\bar{z} \subseteq \bigcap_{\bar{x} \in S} \bar{x} \neq \emptyset.$$

So by Zorn's lemma, we can find a minimal element (in both senses). □

**Lemma.** If $(X, T)$ is a minimal system, then for all $\varepsilon > 0$, there is some $m \in \mathbb{N}$ such that for all $x, y \in X$, we have

$$\min_{|s| \leq m} \rho(T^s x, y) < \varepsilon.$$

*Proof.* Suppose not. Then there exists $\varepsilon > 0$ and points $x_i, y_i \in X$ such that

$$\min_{|s| \leq i} \rho(T^s x_i, y_i) \geq \varepsilon.$$

By compactness, we may pass to subsequences, and assume $x_i \to x$ and $y_i \to y$. By continuity, it follows that

$$\rho(T^s x, y) \geq \varepsilon$$

for all $s \in \mathbb{Z}$. This is a contradiction, since $\bar{x} = X$ by minimality. □

**Theorem** (Topological van der Waerden)**.** Let $(X, T)$ be a dynamical system. Then there exists an $\varepsilon > 0$ such that whenever $r \in \mathbb{N}$, then we can find $x \in X$ and $n \in \mathbb{N}$ such that $\rho(x, T^{in} x) < \varepsilon$ for all $i = 1, \cdots, r$.

*Proof.* Without loss of generality, we may assume $(X, T)$ is a minimal system. We induct on $r$. If $r = 1$, we can just choose $y \in X$, and consider $Ty, T^2 y, \cdots$. Then note that by compactness, we have $s_1, s_2 \in \mathbb{N}$ such that

$$\rho(T^{s_1} y, T^{s_2} y) < \varepsilon,$$

and then take $x = T^{s_1} y$ and $n = s_2 - s_1$.

Now suppose the result is true for $r > 1$, and that we have the result for all $\varepsilon > 0$ and $r - 1$.

**Claim.** For all $\varepsilon > 0$, there is some point $y \in X$ such that there is an $x \in X$ and $n \in \mathbb{N}$ such that

$$\rho(t^{in}x, y) < \varepsilon$$

for all $1 \leq i \leq r$.

Note that this is a different statement from the theorem, because we are not starting at $x$. In fact, this is a triviality. Indeed, let $(x_0, n)$ be as guaranteed by the hypothesis. That is, $\rho(x_0, T^{in}x_0) < \varepsilon$ for all $1 \leq i \leq r$. Then pick $y = x_0$ and $x = T^{-n}x_0$.

The next goal is to show that there is nothing special about this $y$.

**Claim.** For all $\varepsilon > 0$ and for all $z \in X$, there exists $x_z \in X$ and $n \in \mathbb{N}$ for which $\rho(T^{in}x_z, z) < \varepsilon$.

The idea is just to shift the picture so that $y$ gets close to $z$, and see where we send $x$ to. We will use continuity to make sure we don't get too far away.

We choose $m$ as in the previous lemma for $\frac{\varepsilon}{2}$. Since $T^{-m}, T^{-m+1}, \cdots, T^m$ are all uniformly continuous, we can choose $\varepsilon'$ such that $\rho(a, b) < \varepsilon'$ implies $\rho(T^s a, T^s b) < \frac{\varepsilon}{2}$ for all $|s| \leq m$.

Given $z \in X$, we obtain $x$ and $y$ by applying our first claim to $\varepsilon'$. Then we can find $s \in \mathbb{Z}$ with $|s| \leq m$ such that $\rho(T^s y, z) < \frac{\varepsilon}{2}$. Consider $x_z = T^s x$. Then

$$\begin{aligned}
\rho(T^{in}x_z, z) &\leq \rho(T^{in}x_z, T^s y) + \rho(T^s y, z) \\
&\leq \rho(T^s(T^{in}x), T^s y) + \frac{\varepsilon}{2} \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon
\end{aligned}$$

**Claim.** For all $\varepsilon > 0$ and $z \in X$, there exists $x \in X$, $n \in \mathbb{N}$ and $\varepsilon' > 0$ such that $T^{in}(B(x, \varepsilon')) \subseteq B(z, \varepsilon)$ for all $1 \leq i \leq r$.

We choose $\varepsilon'$ by continuity, using the previous claim.

---

The idea is that we repeatedly apply the final claim, and then as we keep moving back, eventually two points will be next to each other.

We pick $z_0 \in X$ and set $\varepsilon_0 = \frac{\varepsilon}{2}$. By the final claim, there exists $z_1 \in X$ and some $n_1 \in \mathbb{N}$ such that $T^{in_1}(B(z_1, \varepsilon_1)) \subseteq B(z_0, \varepsilon_0)$ for some $0 < \varepsilon_1 \leq \varepsilon_0$ and all $1 \leq i \leq r$.

Inductively, we find $z_s \in X$, $n_s \in \mathbb{N}$ and some $0 < \varepsilon_s \leq \varepsilon_{s-1}$ such that

$$T^{in_s}(B(z_s, \varepsilon_s)) \subseteq B(z_{s-1}, \varepsilon_{s-1})$$

for all $1 \leq i \leq r$.

By compactness, $(z_s)$ has a convergent subsequence, and in particular, there exists $i < j \in \mathbb{N}$ such that $\rho(z_i, z_j) < \frac{\varepsilon}{2}$.

Now take $x = z_j$, and

$$n = n_j + n_{j-1} + \cdots + n_{i+1}$$

Then

$$T^{\ell n}(B(x, \varepsilon_j)) \subseteq B(z_i, \varepsilon_i).$$

for all $1 \leq \ell \leq r$. But we know

$$\rho(z_i, z_j) \leq \frac{\varepsilon}{2},$$

and

$$\varepsilon_i \leq \varepsilon_0 \leq \frac{\varepsilon}{2}.$$

So we have

$$\rho(T^{\ell n} x, x) \leq \varepsilon$$

for all $1 \leq \ell \leq r$. □

**Theorem** (Topological Hindman's theorem)**.** Let $(X, T)$ be a dynamical system, and suppose $X = \bar{x}$ for some $x \in X$. Then for any minimal subsystem $Y \subseteq X$, then there is some $y \in Y$ such that $x$ and $y$ are proximal.

**Proposition.** A colouring $c : \mathbb{Z} \to [k]$ is minimal iff $c$ has the bounded gaps property.

*Proof.* Suppose $c$ has the bounded gaps property. We want to show that for all $x \in \bar{c}$, we have $\mathrm{Seq}(x) = \mathrm{Seq}(c)$. We certainly got one containment $\mathrm{Seq}(x) \subseteq \mathrm{Seq}(c)$. Consider any interval $I \subseteq \mathbb{Z}$. We want to show that $c(I) \in \mathrm{Seq}(x)$. By the bounded gaps property, there is some $M$ such that any interval $\mathcal{U} \subseteq \mathbb{Z}$ of length $> M$ satisfies $c(I) \preccurlyeq c(\mathcal{U})$. But since $x \in \bar{c}$, we know that

$$x([0, \cdots, M]) = c([t, t+1, \cdots, t+M])$$

for some $t \in \mathbb{Z}$. So $c(I) \preccurlyeq x([0, \cdots, M])$, and this implies $\mathrm{Seq}(c) \subseteq \mathrm{Seq}(x)$.

For the other direction, suppose $c$ does not have the bounded gaps property. So there is some bad interval $I \subseteq \mathbb{Z}$. Then for any $n \in \mathbb{N}$, there is some $t_n$ such that

$$c(I) \not\preccurlyeq c([t_n - n, t_n - n + 1, \cdots, t_n + n]).$$

Now consider $x_n = \mathcal{L}^{t_n}(c)$. By passing to a subsequence if necessary, we may assume that $x_n \to x$. Clearly, $c(I) \notin \mathrm{Seq}(x)$. So we have found something in $\mathrm{Seq}(c) \setminus \mathrm{Seq}(x)$. So $c \notin \bar{x}$. □

**Theorem** (Hindman's theorem)**.** If $c : \mathbb{N} \to [k]$ is a $k$-colouring, then there exists an infinite $A \subseteq \mathbb{N}$ such that $FS(A)$ is monochromatic.

*Proof.* We extend the colouring $c$ to a colouring of $\mathbb{Z}$ by defining $x : \mathbb{Z} \to [k+1]$ with

$$x(i) = x(-i) = c(i)$$

if $x \neq 0$, and $x(0) = k+1$. Then it suffices to find an infinite an infinite $A \subseteq \mathbb{Z}$ such that $FS(A)$ is monochromatic with respect to $x$.

We apply topological Hindman to $(\bar{x}, \mathcal{L})$ to find a minimal colouring $y$ such that $x$ and $y$ are proximal. Then either

$$\inf_{n \in \mathbb{N}} \rho(\mathcal{L}^n x, \mathcal{L}^n y) = 0 \text{ or } \inf_{n \in \mathbb{N}} \rho(\mathcal{L}^{-n} x, \mathcal{L}^{-n} y) = 0.$$

We wlog it is the first case, i.e. $x$ and $y$ are positively proximal. We now build up this infinite set $A$ we want one by one.

Let the colour of $y(0)$ be red. We shall in fact pick $0 < a_1 < a_2 < \cdots$ inductively so that $x(t) = y(t) = \text{red}$ for all $t \in FS(a_1, a_2, \cdots)$.

By the bounded gaps property of $y$, there exists some $M_1$ such that for any $I \subseteq \mathbb{Z}$ of length $\geq M_1$, then it contains $y([0])$, i.e. $y([0]) \preccurlyeq y(I)$.

Since $x$ and $y$ are positively proximal, there exists $I \subseteq \mathbb{Z}$ such that $|I| \geq M$ and $\min(I) > 0$ such that $x(I) = y(I)$. Then we pick $a_1 \in I$ such that $x(a_1) = y(a_1) = y(0)$.

Now we just have to keep doing this. Suppose we have found $a_1 < \cdots < a_n$ as required. Consider the interval $J = [0, \cdots, a_1 + a_2 + \cdots + a_n]$. Again by the bounded gaps property, there exists some $M_{n+1}$ such that if $I \subseteq \mathbb{Z}$ has $|I| \geq M_{n+1}$, then $y(J) \preccurlyeq y(I)$.

We choose an interval $I \subseteq \mathbb{Z}$ such that

(i) $x(I) = y(I)$

(ii) $|I| \geq M_{n+1}$

(iii) $\min I > \sum_{i=1}^n a_i$.

Then we know that

$$y([0, \cdots, a_1 + \cdots + a_n]) \preccurlyeq y(I).$$

Let $a_{n+1}$ denote by the position at which $y([0, \cdots, a_1 + \cdots + a_n])$ occurs in $y(I)$. It is then immediate that $x(z) = y(z) = \text{red}$ for all $z \in FS(a_1, \cdots, a_{n+1})$, as any element in $FS(a_1, \cdots, a_{n+1})$ is either

- $t'$ for some $t' \in FS(a_1, \ldots, a_n)$;
- $a_{n+1}$; or
- $a_{n+1} + t'$ for some $t' \in FS(a_1, \cdots, a_n)$,

and these are all red by construction.

Then $A = \{a_1, a_2, \cdots\}$ is the required set. □

**Theorem** (Topological Hindman's theorem)**.** If $(X, T)$ is a dynamical system, $X = \bar{x}$ and $Y \subseteq X$ is minimal, then there exists $y \in Y$ such that $x$ and $y$ are proximal.

*Proof.* If $X$ is a compact metric space, then $X^X = \{f : X \to X\}$ is compact under the product topology by Tychonoff. The basic open sets are of the form

$$\{f : X \to Y \mid f(x_i) \in U_i \text{ for } i = 1, \cdots, k\}$$

for some fixed $x_i \in X$ and $U_i \subseteq X$ open.

Now any function $g : X \to X$ can act on $X^X$ in two obvious ways —

- Post-composition $\mathcal{L}_g : X^X \to X^X$ be given by $\mathcal{L}_g(f) = g \circ f$.
- Pre-composition $\mathcal{R}_g : X^X \to X^X$ be given by $R_g(f) = f \circ g$.

The useful thing to observe is that $\mathcal{R}_g$ is always continuous, while $\mathcal{L}_g$ is continuous if $g$ is continuous.

Now if $(X, T)$ is a dynamical system, we let

$$E_T = \mathrm{cl}\{T^s : s \in \mathbb{Z}\} \subseteq X^X.$$

The idea is that we look at what is going on in this space instead. Let's note a few things about this subspace $E_T$.

- $E_T$ is compact, as it is a closed subset of a compact set.
- $f \in E_T$ iff for all $\varepsilon > 0$ and points $x_1, \cdots, x_k \in X$, there exists $s \in \mathbb{Z}$ such that $\rho(f(x_i), T^s(x_i)) < \varepsilon$.
- $E_T$ is closed under composition.

So in fact, $E_T$ is a compact semi-group inside $X^X$. It turns out every proof of Hindman's theorem involves working with a semi-group-like object, and then trying to find an idempotent element. □

**Theorem** (Idempotent theorem)**.** If $E \subseteq X^X$ is a compact semi-group, then there exists $g \in E$ such that $g^2 = g$.

*Proof.* Let $\mathcal{F}$ denote the collection of all compact semi-groups of $E$. Let $A \in \mathcal{F}$ be minimal with respect to inclusion. To see $A$ exists, it suffices (by Zorn) to check that any chain $S$ in $\mathcal{F}$ has a lower bound in $\mathcal{F}$, but this is immediate, since the intersection of nested compact sets is noon-empty.

**Claim.** $Ag = A$ for all $g \in A$.

We first observe that $Ag$ is compact since $\mathcal{R}_g$ is continuous. Also, $Ag$ is a semigroup, since if $f_1 g, f_2 g \in Ag$, then

$$f_1 g f_2 g = (f_1 g f_2) g \in A_g.$$

Finally, since $g \in A$, we have $Ag \subseteq A$. So by minimality, we have $Ag = A$.

Now let

$$B_g = \{f \in A : fg = g\}.$$

**Claim.** $B_g = A$ for all $g \in A$.

Note that $B_g$ is non-empty, because $Ag = A$. Moreover, $B$ is closed, as $B$ is the inverse of $\{g\}$ under $\mathcal{R}_g$. So $B$ is compact. Finally, it is clear that $B$ is a semi-group. So by minimality, we have $B_g = A$.

So pick any $g \in A$, and then $g^2 = g$. □

*Proof of topological Hindman (continued).* We are actually going to work in a slightly smaller compact semigroup than $E_T$. We let $\mathcal{F} \subseteq E_T$ be defined by

$$\mathcal{F} = \{f \in E_T : f(x) \in Y\}.$$

**Claim.** $\mathcal{F} \subseteq X^X$ is a compact semigroup.

Before we prove the claim, we see why it makes sense to consider this $\mathcal{F}$. Suppose the claim is true. Then by applying the idempotent theorem, to get $g \in \mathcal{F}$ such that $g^2 = g$. We now check that $x, g(x)$ are proximal.

Since $g \in \mathcal{F} \subseteq E_T$, we know for all $\varepsilon$, there exists $s \in \mathbb{Z}$ such that

$$\rho(g(x), T^s(x)) < \frac{\varepsilon}{2}, \quad \rho(g(g(x)), T^s(g(x))) < \frac{\varepsilon}{2}.$$

But we are done, since $g(x) = g(g(x))$, and then we conclude from the triangle inequality that

$$\rho(T^s(x), T^s(g(x))) < \varepsilon.$$

So $x$ and $g(x) \in Y$ are proximal.

It now remains to prove the final claim. We first note that $\mathcal{F}$ is non-empty. Indeed, pick any $y \in Y$. Since $X = \bar{x}$, there exists some sequence $T^{n_i}(x) \to y$. Now by compactness of $E_T$, we can find some $f \in E_T$ such that $(T^{n_i})_{i \geq 0}$ cluster at $f$, i.e. every open neighbourhood of $f$ contains infinitely many $T^{n_i}$. Then since $X$ is Hausdorff, it follows that $f(x) = y$. So $f \in \mathcal{F}$.

To show $\mathcal{F}$ is compact, it suffices to show that it is closed. But since $Y = \bar{y}$ for all $y \in Y$ by minimality, we know $Y$ is in particular closed, and so $\mathcal{F}$ is closed in the product topology.

Finally, we have to show that $\mathcal{F}$ is closed under composition, so that it is a semi-group. To show this, it suffices to show that any map $f \in E_T$ sends $Y$ to $Y$. Indeed, by minimality of $Y$, this is true for $T^s$ for $s \in \mathbb{Z}$, and then we are done since $Y$ is closed. □

# 5 Sums and products*

**Theorem** (Moreira, 2016)**.** If $\mathbb{N}$ is $k$-coloured, then there exists infinitely many $x, y$ such that $\{x, x+y, xy\}$ is monochromatic.

**Proposition.** If $A$ is piecewise syndetic and $A = X \cup Y$, then one of $X$ and $Y$ is piecewise syndetic.

*Proof.* We fix $b > 0$ and the sequence of intervals $I_1, I_2, \cdots$ with $|I_n| \geq n$ such that $A$ has gaps $\leq b$ in each $I_n$. We may, of course, wlog $I_n$ are disjoint.

Each element in $I_n$ is either in $X$ or not. We let $t_n$ denote the largest number of consecutive things in $I_n$ that are in $X$. If $t_n$ is unbounded, then $X$ is piecewise syndetic. If $t_n \leq K$, then $Y$ is piecewise syndetic, with gaps bounded by $(K+1)b$. □

**Proposition.** Let $A \subseteq \mathbb{N}$ be piecewise syndetic. Then for all $m \in \mathbb{N}$, there exists $d \in \mathbb{N}$ such that the set

$$A^* = \{x \in \mathbb{N} : x, x+d, \cdots, x+md \in A\}.$$

is piecewise syndetic.

*Proof.* Let's in fact assume that $A$ is syndetic, with gaps bounded by $b$. The proof for a piecewise syndetic set is similar, but we have to pass to longer and longer intervals, and then piece them back together.

We want to apply van der Waerden's theorem. By definition,

$$\mathbb{N} = A \cup (A+1) \cup \cdots \cup (A+b).$$

Let $c : \mathbb{N} \to \{0, \cdots, b\}$ by given by

$$c(x) = \min\{i : x \in A + i\}.$$

Then by van der Waerden, we can find some $a_1, d_1$ such that

$$a_1, a_1 + d_1, \cdots, a_1 + md_1 \in A + i.$$

Of course, this is equivalent to

$$(a_1 - i), (a_1 + i) + d_1, \cdots, (a_1 - i) + md_1 \in A.$$

So we wlog $i = 0$.

But van der Waerden doesn't require the whole set of $\mathbb{N}$. We can split $\mathbb{N}$ into huge blocks of size $W(m+1, b+1)$, and then we run the same argument in each of these blocks. So we find $(a_1, d_1), (a_2, d_2), \cdots$ such that

$$a_i, a_i + d_i, \cdots, a_i + md_i \in A.$$

Moreover, $\tilde{A} = \{a_1, a_2, \cdots\}$ is syndetic with gaps bounded by $W(m+1, b+1)$. But we need a fixed $d$ that works for everything.

We now observe that by construction, we must have $d_i \leq W(m+1, b+1)$ for all $i$. So this induces a finite colouring on the $a_i$, where the colour of $a_i$ is just $d_i$. Now we are done by the previous proposition, which tells us one of the partitions must be piecewise syndetic. □

*Proof of theorem.* Let $\mathbb{N} = C_1 \cup \cdots \cup C_r$. We build the following sequences:

(i) $y_1, y_2, \cdots \in \mathbb{N}$;

(ii) $B_0, B_1, \cdots$ and $D_1, D_2, \cdots$ which are piecewise syndetic sets;

(iii) $t_0, t_1, \cdots \in [r]$ which are colours.

We will pick these such that $B_i \subseteq C_{t_i}$.

We first observe that one of the colour classes $C_i$ is piecewise syndetic. We pick $t_0 \in [r]$ be such that $C_{t_0}$ is piecewise syndetic, and pick $B_0 = C_{t_0}$. We want to apply the previous proposition to obtain $D_0$ from $B_0$. We pick $m = 2$, and then we can find a $y_1$ such that

$$D_1 = \{z \mid z, z + y_1 \in B_0\}$$

is piecewise syndetic.

We now want to get $B_1$. We notice that $y_1 D_1$ is also piecewise syndetic, just with a bigger gap. So there is a colour class $t_1 \in [r]$ such that $y_1 D_1 \cap C_{t_1}$ is piecewise syndetic, and we let $B_1 = y_1 D_1 \cap C_{t_1}$.

In general, having constructed $y_1, \cdots, y_{i-1}$; $B_0, \cdots, B_{i-1}$; $D_1, \cdots, D_{i-1}$ and $t_0, \ldots, t_{i-1}$, we proceed in the following magical way:

We apply the previous proposition with the really large $m = y_1^2 y_2^2 \cdots y_{i-1}^2$ to find $y_i \in \mathbb{N}$ such that the set

$$D_i = \{z \mid z, z + y_i, \cdots, z + (y_1^2 y_2^2 \cdots y_{i-1}^2) y_i \in B_{i-1}\}$$

is piecewise syndetic. The main thing is that we are not using all points in this progression, but are just using some important points in there. In particular, we use the fact that

$$z, z + y_i, z + (y_j^2 \cdots y_{i-1}^2) y_i \in B_{i-1}$$

for all $1 \leq j \leq i - 1$. It turns out these squares are important.

We have now picked $y_i$ and $D_i$, but we still have to pick $B_i$. But this is just the same as what we did. We know that $y_i D_i$ is piecewise syndetic. So we know there is some $t_i \in [r]$ such that $B_i = y_i D_i \cap C_{t_i}$ is piecewise syndetic.

Let's write down some properties of these $B_i$'s. Clearly, $B_i \subseteq C_{t_i}$, but we can say much more. We know that

$$B_i \subseteq y_i B_{i-1} \subseteq \cdots \subseteq y_i y_{i-1} \cdots y_{j+1} B_j.$$

Of course, there exists some $t_j, t_i$ with $j < i$ such that $t_j = t_i$. We set

$$y = y_i y_{i-1} \cdots y_{j+1}.$$

Let's choose any $z \in B_i$. Then we can find some $x \in B_j$ such that

$$z = xy,$$

We want to check that this gives us what we want. By construction, we have

$$x \in B_j \subseteq C_{t_i}, \quad xy = z \in B_i \subseteq C_{t_i}.$$

So they have the same colour.

How do we figure out the colour of $x + y$? Consider

$$y(x+y) = z + y^2 \in y_i D_i + y^2.$$

By definition, if $a \in D_i$, then we know

$$a + (y_{j+1}^2 \cdots y_{i-1}^2) y_i \in B_{i-1}.$$

So we have

$$D_i \subseteq B_{i-1} - (y_{j+1}^2 \cdots y_{i+1}^2) y_i \subseteq y_{j+1} \cdots y_{i-1} B_j - (y_{j+1}^2 \cdots y_{i-1}^2) y_i$$

So it follows that

$$y(x+y) \subseteq y_i (y_{i-1} \cdots y_{j+1}) B_j - y^2 + y^2 = y B_j.$$

So $x + y \in B_j \subseteq C_{t_j} = C_{t_i}$. So $\{x, x+y, xy\}$ have the same colour. $\square$